INVESTIGAÇÃO DA INFLUÊNCIA DO TAMANHO DE ROIs NA SEGMENTAÇÃO AUTOMÁTICA DE NÓDULOS EM IMAGENS DE SIMULADORES MAMOGRÁFICOS

K. D. Marcomini* e H. Schiabel*

*Departamento de Engenharia Elétrica – EESC/Universidade de São Paulo, São Carlos, Brasil
e-mail: karem.dm@usp.br

Resumo: As interpretações falsas-negativas são graves problemas no diagnóstico médico. Para reduzir esses casos, métodos computacionais vêm sendo desenvolvidos e aperfeiçoados a fim de aumentar a sensibilidade diagnóstica e auxiliar na detecção precoce do câncer de mama. Porém, esses sistemas sofrem influência direta ou indiretamente do usuário. Nesse sentido, avaliamos como a falta de padronização no recorte de regiões de interesse pode influenciar na etapa de detecção computadorizada. Para isso, utilizamos 100 lesões representadas por simuladores mamográficos expostos a raios X. Cada lesão foi recortada três vezes, variando a quantidade de tecido adjacente. As regiões de interesse foram segmentadas pelas técnicas de contorno ativo e crescimento de região. O resultado mais preciso foi alcançado com o recorte realizado mais próximo possível da lesão em ambos os métodos. As técnicas de contorno ativo e crescimento de região proporcionaram sensibilidade de 97,18% e 91,56%; e acurácia de 96,70% e 95,82%, respectivamente. No contorno ativo, foi possível atingir resultados similares nos outros dois recortes, porém o processamento torna-se mais lento devido ao aumento de iterações. Dessa forma, podemos concluir que a quantidade de tecido adjacente a lesão influência diretamente nos parâmetros iniciais dos segmentadores, sendo necessário repassar ao radiologista o padrão de como o recorte deve ser realizado.
Palavras-chave: câncer de mama, mamografia, região de interesse, segmentação.

Abstract: *The false-negative cases interpretations represent serious problems in medical diagnosis. To reduce the number of these cases and increase the diagnostic sensibility, it have been developed computational tools to aiding the early detection of breast cancer. However, the user can influence directly or indirectly these systems. In this context, we evaluate how the lack of standardization in cutting image regions of interest (ROIs) can affect the computed detection step. We utilize 100 mammographic images lesions represented by simulators exposed to X-rays. A radiologist cropped each lesion three times varying the amount of surrounding tissue. Each lesion was cropped three times, varying the amount of surrounding tissue. The regions of interest have been segmented by active contour and region growing techniques. The most accurate result was obtained with the cutting performed as close as of the lesion in both methods. Active contour and region growing techniques provided sensitivity 97.18% and 91.56%; and accuracy of 96.70% and 95.82%, respectively. In the active contour, similar results could be achieved in the other two cuttings, but the processing became slower due to the increase of iterations. Thus, we can conclude that the amount of tissue surrounding the lesion directly influence the initial parameters of the proposed models, being necessary to transmit the radiologist a pattern to perform the cutting.*
Keywords: *breast cancer, mammography, region of interest, segmentation.*

Introdução

A mamografia é o melhor método de detecção precoce do câncer de mama, sendo a sua interpretação um desafio para o especialista. Cerca de 10 a 30% das lesões mamárias são perdidas nos exames de rotina, devido a limitações próprias dos observadores humanos [1].

Com o avanço do processamento de imagem digital, o reconhecimento de padrões e o uso de inteligência artificial, os radiologistas têm a oportunidade de melhorar seu diagnóstico com o auxílio de sistemas de computador. A detecção auxiliada por computador (*Computer Aided-Detection* – CAD) é uma tecnologia que tem sido implementada em alguns serviços de mamografia, com o intuito de promover dupla leitura. Um sistema CAD é útil em situações em que exista um grande número de especialistas com opiniões distintas (variabilidade interobservador), falta de observadores treinados, ou na impossibilidade de realizar dupla leitura com dois ou mais radiologistas. Estudos clínicos têm demonstrado que o CAD aumenta a sensibilidade de detecção do câncer de mama por radiologistas em até 20% [1].

Nesse âmbito, os sistemas CAD estão sujeitos a diferentes tipos de variabilidade produzida pelo usuário. Uma delas está relacionada à determinação da região de interesse (ROI – do inglês: *Region of Interest*). Assim, a quantidade de tecido adjacente presente na ROI pode exercer influência direta sobre as operações numéricas do segmentador, afetando a delimitação precisa da lesão, principalmente se tratando do processamento digital automático, cujos parâmetros iniciais não são readequados conforme a variação nas características das imagens de entrada. Erros durante a etapa de

segmentação podem comprometer a classificação dos dados, fase essa totalmente dependente do resultado adquirido pela segmentação [2].

No presente trabalho propomos uma análise do efeito que a quantidade de tecido adjacente presente na imagem pode ter no processo de segmentação automática de mamografias provenientes de simuladores da mama.

Materiais e métodos

Para o desenvolvimento desse estudo foram adotadas as imagens produzidas a partir da exposição do simulador mamográfico desenvolvido por [3] aos seguintes mamógrafos: *Lorad MIV*, instalado no Centro Integrado de Diagnóstico por Imagem da Santa Casa de São Carlos (SP); e *GE Senographe DMR*, do Hospital das Clínicas de Ribeirão Preto da FMRP/USP. Na Figura 1 são apresentadas as imagens do simulador mamográfico utilizado nessa pesquisa. Já a Figura 2 exibe alguns modelos das imagens mamográficas produzidas.

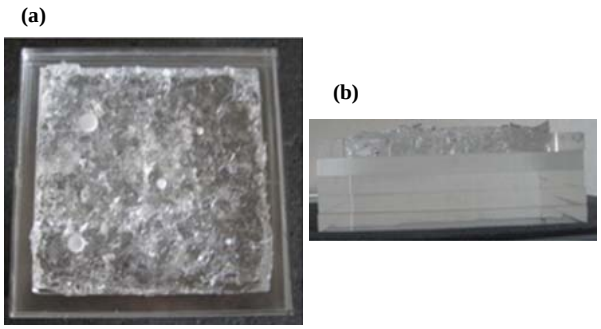

Figura 1: Simulador mamográfico. (a) vista superior; (b) vista lateral.

Figura 2: Imagens radiográficas provenientes da exposição do simulador à raios-X.

No sentido de avaliar quanto a quantidade de tecido adjacente à lesão influencia no processo de segmentação automática de lesões em imagens mamográficas de simuladores mamários, contamos com o apoio de um radiologista experiente para realizar o recorte das ROIs. Foram, então, realizados três cortes diferenciados para a mesma lesão. O primeiro deles corresponde a inclusão mínima de área circunjacente ao achado suspeito; já o segundo corte foi realizado a partir das medidas (área total da ROI – quantidade de *pixels* presentes) do primeiro recorte, o qual teve sua proporção ampliada em 50%; e por fim, o terceiro compreendeu o dobro da área do recorte 1. A Figura 3 apresenta um modelo de cada recorte para o mesmo achado mamográfico.

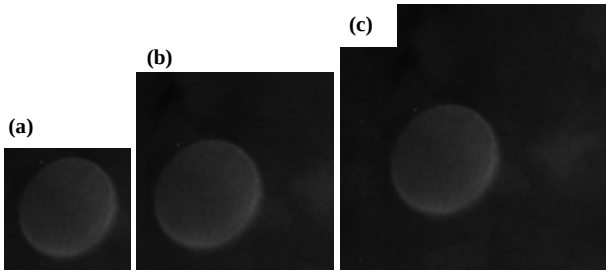

Figura 3: Recortes: (a) próximo à lesão – Recorte 1; (b) quantidade intermediária de tecido adjacente – Recorte 2; (c) maior abrangência de tecido – Recorte 3.

O banco de imagens desse trabalho é composto por 100 lesões distintas, sendo que a mesma lesão foi recortada 3 vezes, totalizando, portanto, 300 ROIs – 100 ROIs para cada tipo de recorte.

Adquiridas as imagens de teste, empregamos duas técnicas de segmentação: contorno ativo e crescimento de região.

O contorno ativo é caracterizado por tentar ajustar uma curva ou contorno inicial sobre a imagem. A movimentação da curva ocorre de forma a tentar minimizar a energia funcional que é dada pela combinação das energias da curva e da imagem, ou seja, são forças que regulam a movimentação constante da curva que se desloca buscando encontrar a borda do objeto e se ajusta aos níveis mínimos de energia [4-5]. Trata-se de um processo interativo (curva inicial é definida pelo usuário) e iterativo (deve-se determinar um "tempo" ou número máximo de iterações em busca do ponto mínimo de energia). Mediante essas variáveis, propomos a variação na razão de aspecto do contorno inicial com forma retangular e dimensões de 40, 60 e 80% na quantidade de *pixels* presente na altura e largura da ROI, estando localizado na região central da imagem. Quanto ao número máximo de iterações adotamos os valores de 200 e 500 ciclos, verificando ainda o resultado que proporcionasse maior precisão e gerasse menor custo computacional.

O crescimento de região realiza a segmentação de uma imagem em relação a um conjunto de pontos, conhecidos como sementes. A partir daí, a imagem é dividida em sub-regiões na qual irão ser agregados os *pixels* similares. A união dessas regiões depende da análise de similaridade, em que os *pixels* são examinados baseando-se em um *threshold*. Sendo assim, posicionamos a semente no *pixel* central da imagem e o valor utilizado para o *threshold* foi dado pelo método de otsu [6].

Para avaliar quantitativamente os dados obtidos contamos novamente com a ajuda de uma radiologista para delinear manualmente o contorno da lesão. Esse contorno foi delineado de maneira que fosse igual para todas as ROIs.

As ROIs demarcadas manualmente foram comparadas com as detectadas automaticamente pelo segmentador, e, dessa comparação, derivaram três medidas quantitativas: Medida Combinada (CM –

Combination Measure), Sensibilidade e Acurácia [7-8].

Resultados

Para analisar a influência que a quantidade de tecido adjacente à lesão exerce sobre os algoritmos propostos, comparamos quantitativamente a segmentação automática com a delimitada manualmente pelo radiologista com base em três medidas.

Nas Tabelas de 1 a 3 temos os dados referentes a segmentação por contorno ativo, apresentando as mudanças em alguns parâmetros de entrada. Já na Tabela 4 são apresentados os resultados provenientes da segmentação por crescimento de região.

Tabela 1: Medidas de avaliação da segmentação por contorno ativo aplicadas no Recorte 1.

	200 Ciclos			**500 Ciclos**		
	Retângulo 40%	Retângulo 60%	Retângulo 80%	Retângulo 40%	Retângulo 60%	Retângulo 80%
CM	82,48	94,61	94,60	89,97	94,73	94,67
Sensibilidade	74,69	95,80	97,18	83,30	95,86	96,19
Acurácia	86,95	96,73	96,70	90,16	96,81	96,77

Tabela 2: Medidas de avaliação da segmentação por contorno ativo aplicadas no Recorte 2.

	200 Ciclos			**500 Ciclos**		
	Retângulo 40%	Retângulo 60%	Retângulo 80%	Retângulo 40%	Retângulo 60%	Retângulo 80%
CM	88,46	89,27	66,43	92,89	93,40	79,22
Sensibilidade	85,53	97,43	99,52	91,26	96,11	98,39
Acurácia	95,50	96,26	73,37	96,35	97,98	81,18

Tabela 3: Medidas de avaliação da segmentação por contorno ativo aplicadas no Recorte 3.

	200 Ciclos			**500 Ciclos**		
	Retângulo 40%	Retângulo 60%	Retângulo 80%	Retângulo 40%	Retângulo 60%	Retângulo 80%
CM	88,35	71,97	47,03	91,06	90,10	53,78
Sensibilidade	86,66	97,44	99,67	87,31	95,61	99,50
Acurácia	95,52	91,94	49,66	96,07	97,18	53,90

Tabela 4: Medidas de avaliação da segmentação por crescimento de região aplicadas nos três recortes.

	Recorte 1	Recorte 2	Recorte 3
CM	93,30	89,70	84,91
Sensibilidade	91,56	90,67	91,19
Acurácia	95,82	93,49	86,86

Discussão

Antes de realizar o recorte de uma ROI deve-se repassar ao especialista as condições que o mesmo deve ser feito, pois a quantidade de tecido adjacente à lesão pode influenciar as operações numéricas do segmentador.

No caso do contorno ativo, por exemplo, quando o contorno inicial é posicionado próximo aos pontos de borda o risco de convergir para locais errados é menor e torna o processo mais rápido.

O fator acima é evidente no recorte 1, em que foram obtidos os melhores resultados quando o retângulo representando o contorno inicial é maior, pois necessita de menos iterações, se aproximando mais da borda da lesão, que são os retângulos com razão de aspecto de 60 e 80% do tamanho da ROI. Já para a proporção de 40% é necessária maior quantidade de ciclos para garantir a precisão da técnica.

Já no caso dos recortes 2 e 3, quando é admitido acrescentar maior quantidade de tecido adjacente, a lesão pode não estar mais localizada na região central da imagem, o que torna necessário quantidades maiores de iterações (valores mais precisos com 500 ciclos) e menor retângulo inicial (razão de aspecto de 40 e 60%).

A Figura 4 apresenta um exemplo de como a segmentação por contorno ativo pode ocorrer de maneira

imprecisa no caso do recorte do especialista se mostrar sem um critério pré-definido. Nessa figura são expostos os três recortes a uma mesma condição inicial, cujos parâmetros são: retângulo inicial com proporção de 80% e 500 iterações.

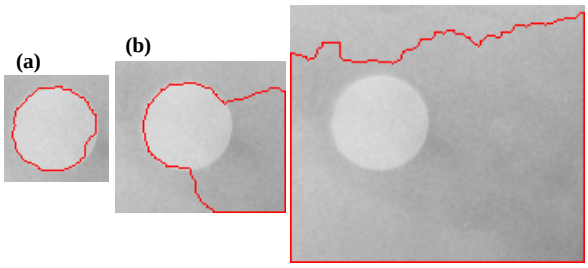

Figura 4: Segmentação das ROIs: (a) próximo à lesão (b) quantidade intermediária de tecido adjacente (c) maior abrangência de tecido.

A técnica de crescimento de região alcançou maior precisão no recorte 1 considerando as três medidas adotadas, pois ocorre menor influência da intensidade de fundo nas operações numéricas e, principalmente, sobre o valor do *threshold*.

Conclusão

A técnica de segmentação por contorno ativo possui um elevado custo computacional, e tente a aumentar ainda mais quando torna-se necessário mais iterações para a convergência dos dados. Nesse sentido, quanto menor for o tamanho do recorte, mais rápido se dá o processo.

Para ambas as técnicas de segmentação selecionadas, considera-se que o melhor resultado foi obtido com o recorte 1, o qual proporcionou taxas mais precisas na comparação com a delimitação manual do radiologista. No caso do contorno ativo, é importante ressaltar que o retângulo inicial deve ter dimensão igual a superior a 60% sendo necessárias apenas 200 iterações.

Dessa forma, para que o algoritmo obtenha maior precisão possível na segmentação de lesões mamárias, é necessário que haja uma uniformização na maneira que o recorte é realizado, pois se o sistema estiver habilitado para receber regiões com a mínima quantidade de tecido adjacente, caso o usuário apresente uma ROI com quantidade significativamente maior ou menor de tecido circunjacente à lesão o sistema pode segmenta-la erroneamente. Assim, é importante que as informações sejam repassadas ao profissional que irá utilizar o *software*.

Agradecimentos

À FAPESP pelo apoio financeiro.

Referências

[1] Calas MJG, Gutfilen B, Pereira WCA, “CAD e mamografia: por que usar esta ferramenta?”. *Radiologia Brasileira*, v. 45, p. 46-52, 2012.

[2] Gruszauskas NP, Drukker K, Giger M, Sennett CA, Pesce LL. “Performance of breast ultrasound computer aided-diagnosis: Dependence on image selection”. *Academic Radiology*, v. 15, n. 10, p. 1234-1245, out. 2008.

[3] Stemberg D. “*Desenvolvimento de um simulador mamográfico para avaliação de esquemas CAD*”. 82 f. Dissertação de mestrado, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2010.

[4] Sandhu R, Georgiou T, Tannembaum A. “A new distribution metric for image segmentation”. Medical Imaging 2008: Image Processing. *Proc. SPIE 2008*, Marseille, França, v. 6914, p. 691404-1–691404-9, fev. 2008.

[5] Wirth MA, Stapinski A. “Segmentation of the breast region in mammograms using active contours”. *Proc. SPIE 2003*, Orlando (FL), USA, v. 5150, p. 1995-2006, fev. 2003.

[6] Nagi J, Kareem SA, Nagi F, Ahmed SK. “Automated breast profile segmentation for ROI detection using digital mammograms”. *In Biomedical Engineering and Sciences (IECBES)*, p. 87-92, 2010.

[7] Boaventura IAG. “*Números fuzzy em processamento de imagens digitais e suas aplicações na detecção de borda*”. 2010. 109 f. Tese de Doutorado, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2010.

[8] Pei C, Wang C, Xu S. “Segmentation of the breast region in mammograms using marker-controlled watershed transform”. *In: 2nd International Conference on Information Science and Engineering*, Hangzhou, China, Proceedings of ICISE’2010, p. 2371-2374, dez. 2010.